Supplementary Figure E1. Correlations of candidate serum markers with blood counts**.** As blood count improvements after IST could be a confounding parameter in biomarker selection, Pearson correlation analysis was performed using a web-based tool. Pairwise complete observation was chosen for handling NA values, and hierarchical clustering was applied. Correlation with a $p > 0.05$ are not displayed. Blood counts are shown in red: ALC, absolute lymphocyte count; ARC, absolute reticulocyte count; ANC, absolute neutrophil count; AMC, absolute monocyte count; Hb, hemoglobin; PLT, platelet count.

Supplementary Figure E2. Receiver-operating curve analysis for specificity and sensitivity and diagnostic power of selected markers for AA diagnosis. (**A**) Sensitivity and specificity were calculated by a receiver-operating curve analysis, using the healthy control group as a reference. AUC, area under the curve; CI, confidential interval. A combined protein marker panel was used to predict the diagnosis of SAA in the verification (**B**) or validation (**C**) cohorts using the HC group as reference. AUC, area under the curve; CI, confidence interval; $*p < 0.05$; $****p < 0.0001$.

Supplementary Figure E3. Serum levels of candidate biomarkers in the verification cohort and their correlations with blood counts. (**A**) Protein levels were measured by Luminex, and data for HGF and SELL are shown as minimum to maximum for each group of patients. (**B**) Correlation matrix used to build correlograms in Fig. 4. By using all four markers, a generalized linear model analysis was calculated on patients at 1-year follow-up and used to predict responsiveness to IST at baseline (**C**), or after 6 months of treatment (**D**). CR, complete responders; PR, partial responders; mPR, minimal partial responders; NR, non-responders; IST, immunosuppressive therapies; ALC, absolute lymphocyte count; PLT, platelets; ARC, absolute reticulocyte count; Hb, hemoglobin levels; ANC, absolute neutrophil count; AMC, absolute monocyte count; AUC,

area under the curve.

Supplementary Figure E4. Common proteins in serum and plasma samples from aplastic anemia (AA) patients before immunosuppressive therapies (IST). (**A**) Unpaired t-test was performed in serum or plasma proteins between healthy controls (HC) and AA patients before therapy. From this analysis, 478 or 16 proteins were present in serum or plasma of AA patients, respectively, and 27 were common in both signatures. Using this group of 27 markers, a heatmap with hierarchical clustering (**B**) and Pearson correlations with blood counts (**C**, highlighted in red) are displayed. Correlations with $p > 0.05$ are shown as blank squares. WBC, white blood count; ANC, absolute neutrophil count; Hb, hemoglobin levels; PLT, platelet count; ARC, absolute reticulocyte count; AMC, absolute monocyte count; ALC, absolute lymphocyte count.

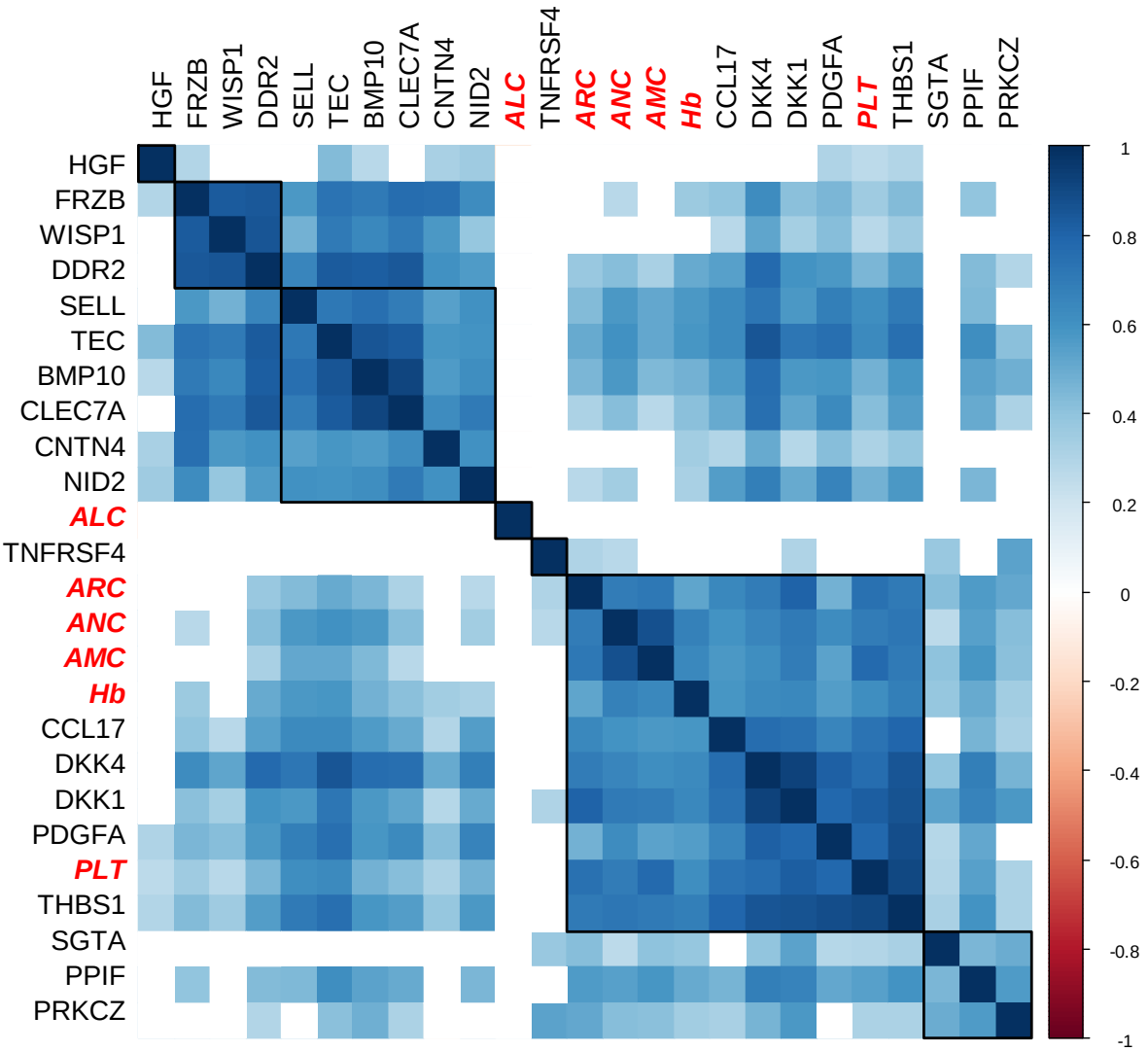
Supplementary Figure E5. Plasma proteomic signature in patients after immunosuppressive therapies (IST). (**A**) After unpaired t-test between HC and patient groups, proteins present in each population after IST were grouped and are shown using Venn diagrams. PCA was performed using 24 proteins present only in CR patients (left panel), or 60 proteins in NR subjects (right panel). Protein pathway analysis was carried out using these two groups of proteins, and top 20 pathways in CR (**B**) or 16 pathways in NR (**C**) are shown.

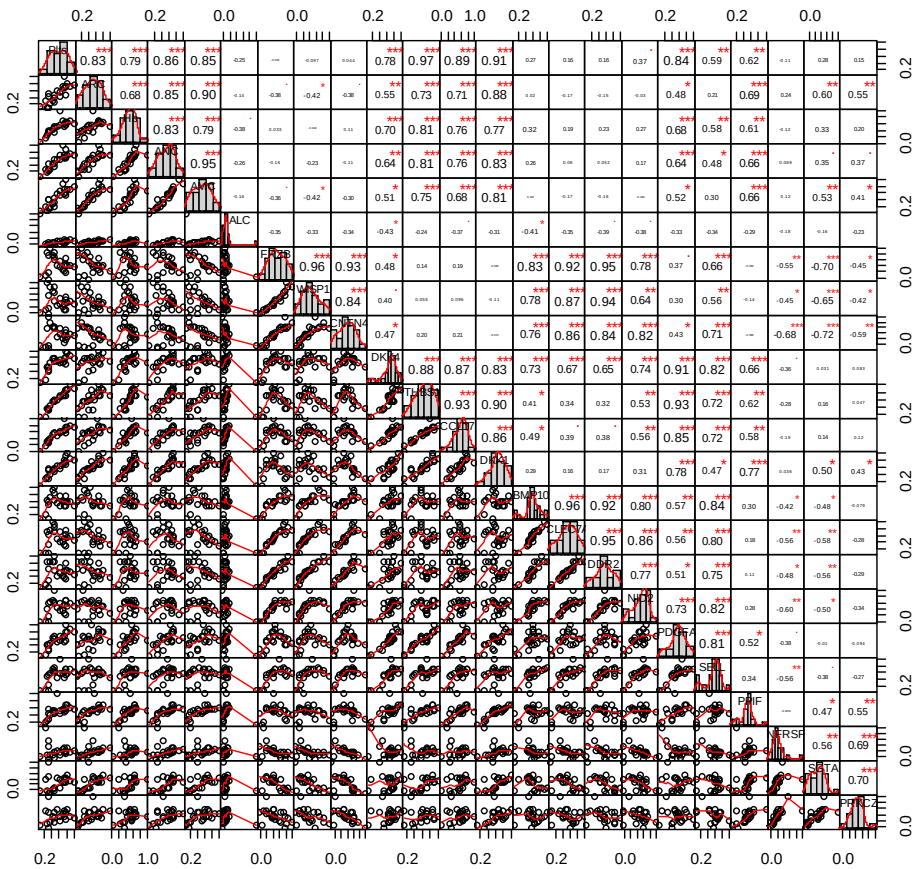
Supplementary Figure E6. Serum and plasma protein pathway analysis in complete responder (CR) patients after IST. Unpaired t-test in serum and plasma proteins was performed between complete responders (CR) before and after therapy as described in Figure 2. Proteins higher in CR after immunosuppressive therapies (IST) were grouped and interpolated using Venn diagrams. Using proteins present both in serum and plasma (87), protein pathway analysis was performed, and top 50 pathways are displayed. ATG, anti-thymocyte globulin; CsA, cyclosporine A; EPAG, eltrombopag.

Supplementary Figure E7. Validation of cytokine signatures in plasma. Heatmap with hierarchical clustering (**A**) and Pearson correlations with blood counts (**B**, highlighted in red) are shown using cytokines already reported in aplastic anemia (AA). (**C**) Data are also reported as mean+SD in healthy controls (HC) and AA patients before or after immunosuppressive therapies (IST). Cytokine expression was compared between AA patients and HC by unpaired t-test. (**D**) Plasma levels of TPO and free c-MPL are reported as mean+SD in healthy controls (HC) and complete (CR) or non-responder (NR) patients before or after immunosuppressive therapies (IST). Groups were compared to the HC by one-way ANOVA with Dunnett's multiple comparisons test. WBC, white blood count; ANC, absolute neutrophil count; Hb, hemoglobin levels; PLT, platelet count; ARC, absolute reticulocyte count; AMC, absolute monocyte count; ALC, absolute lymphocyte count; RFU, relative fluorescence unit. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$, ****, $p < 0.0001$.

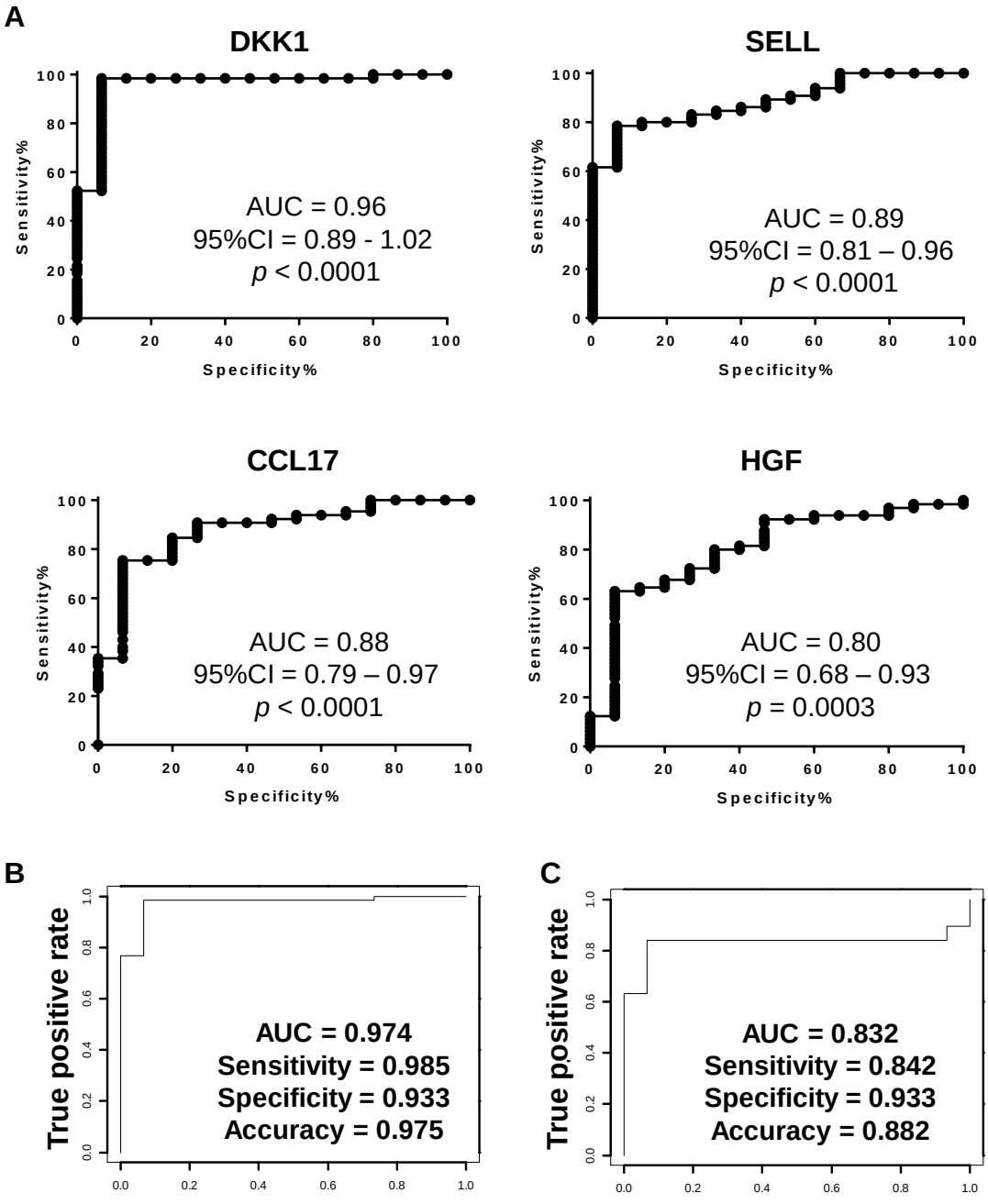
Supplementary Figure E8. Multiple linear regression analysis between disease severity and combined markers. Multiple linear regression analysis between disease severity and validated biomarkers (DKK1, CCL17, SELL, and HGF) was performed by RStudio. Disease severity was defined as the combination of absolute neutrophil count (ANC), absolute reticulocyte count (ARC), and platelet (Plts) count. Linear regression was assessed in plasma (A) and serum (B) samples using corresponding discovery sets. For each analysis, Residuals vs Fitted plot for non-linear patterns, Scale-location plot for variance visualization, Normal Q-Q plot for normal distribution, and Residuals vs Leverage plot for influential outlier identification are shown. Statistics is reported on the right of each panel. *, $p < 0.05$.

Supplementary Figure 1. Giudice V. et al.

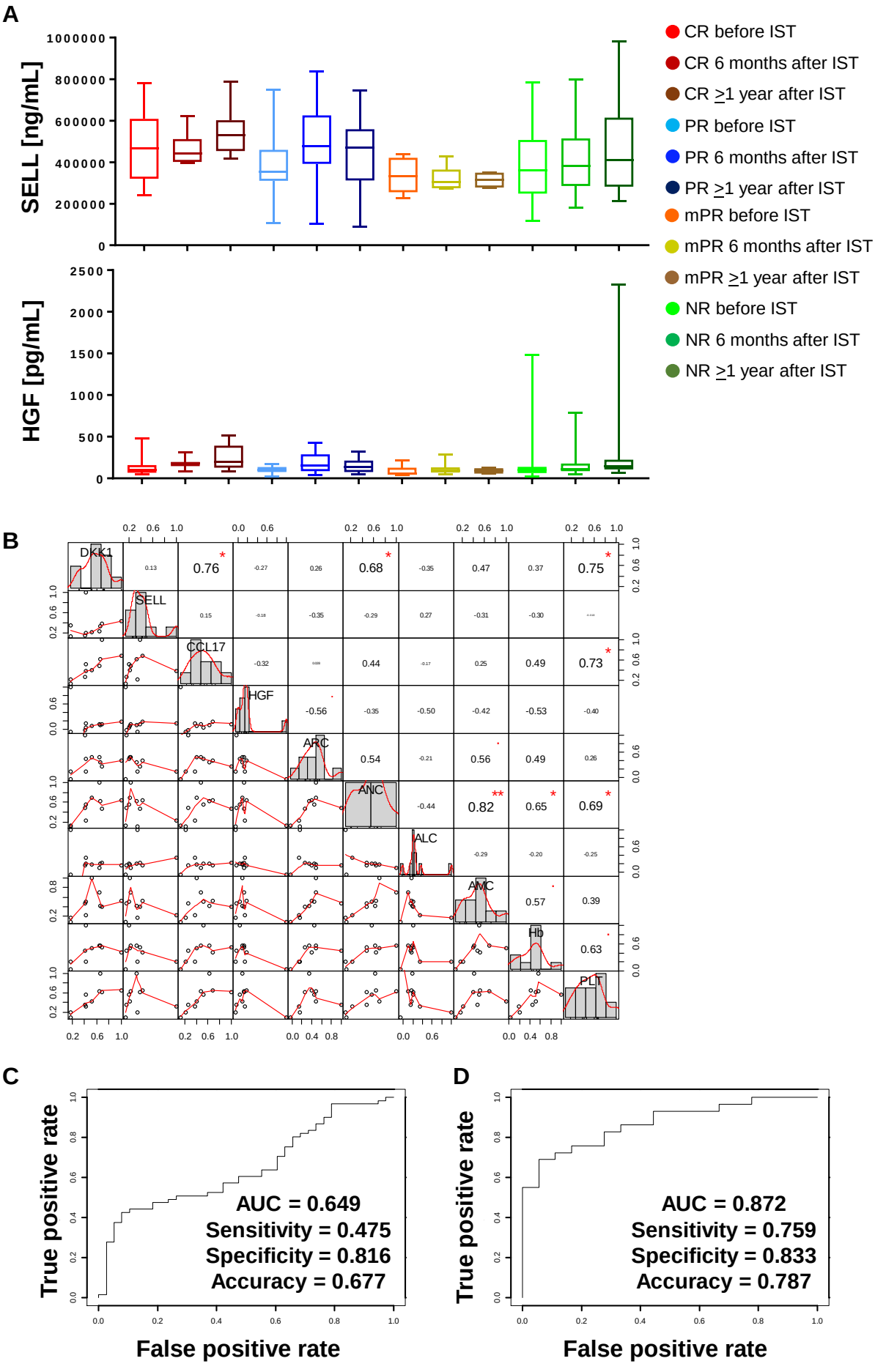


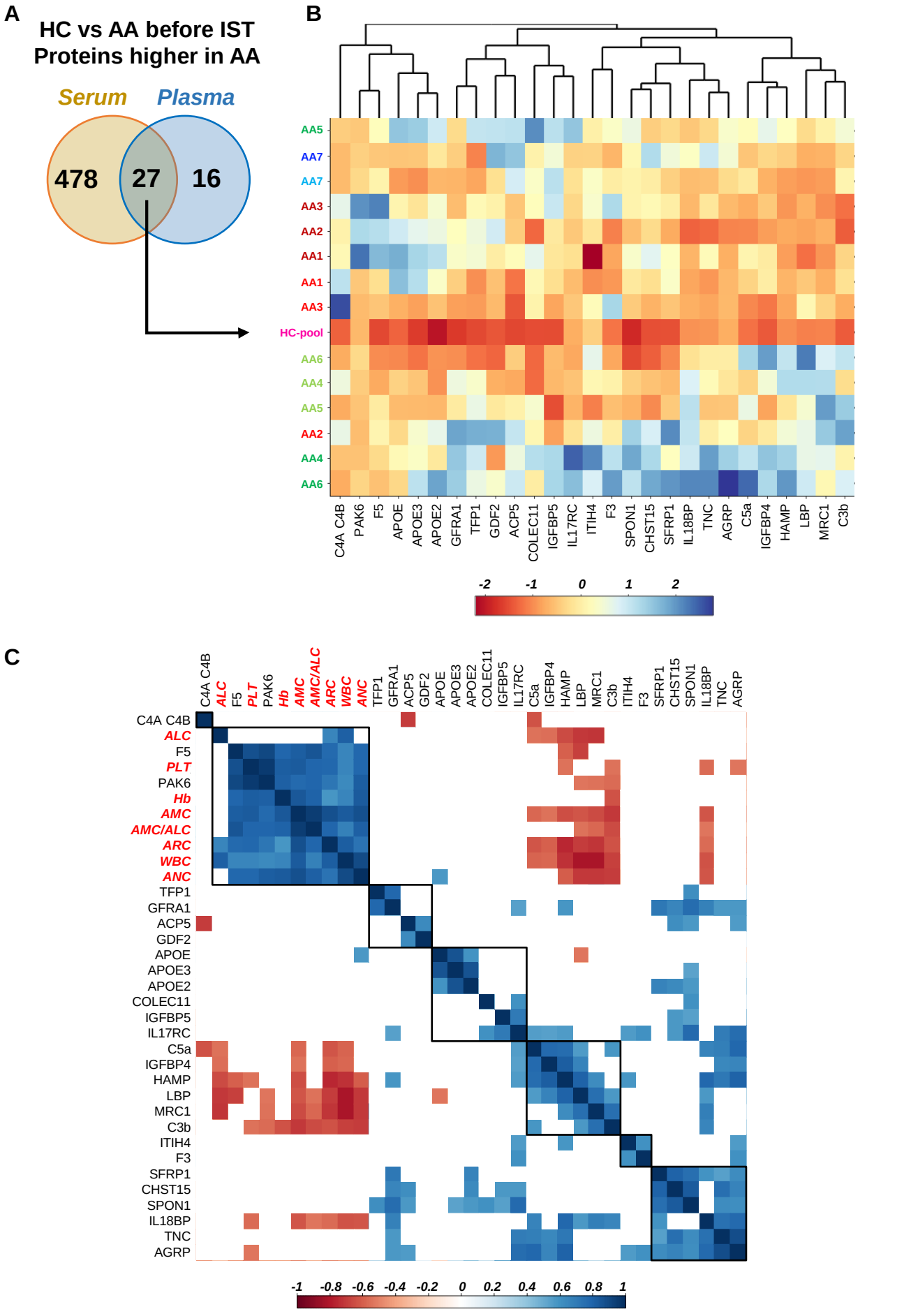


Supplementary Figure E2. Giudice V. et al.

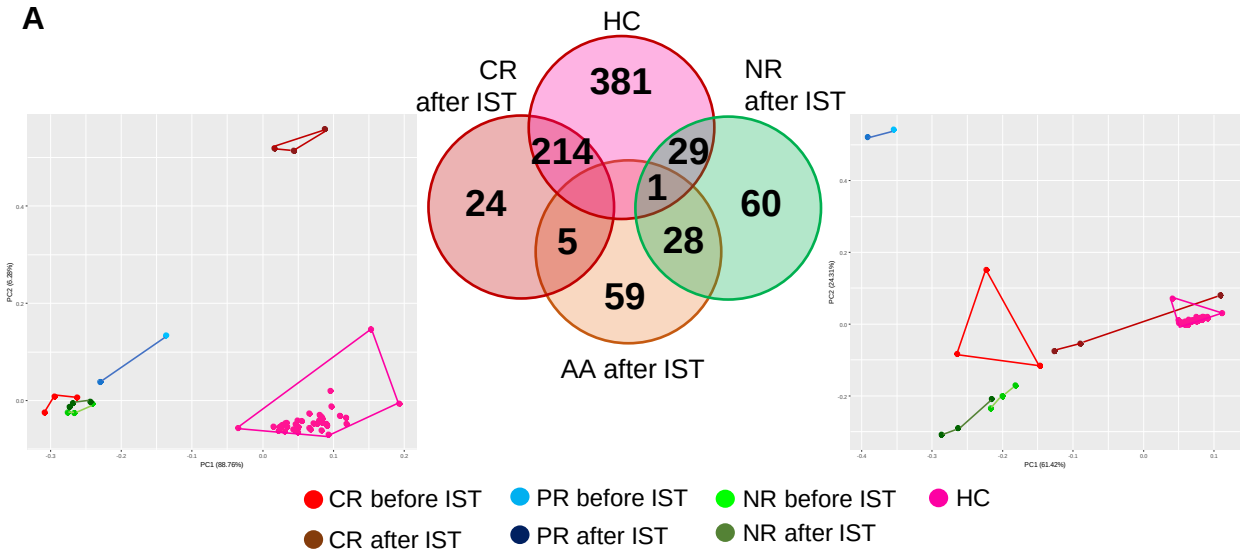


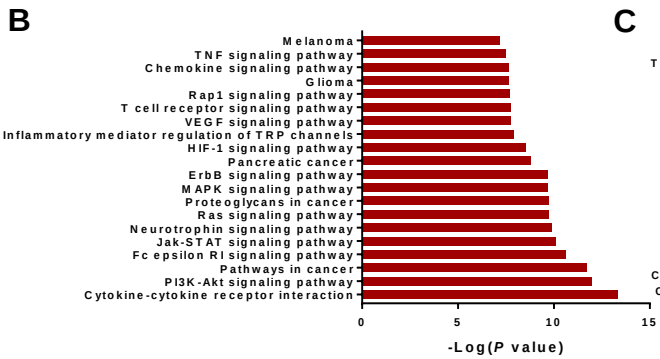
Supplementary Figure E3. Giudice V. et al.

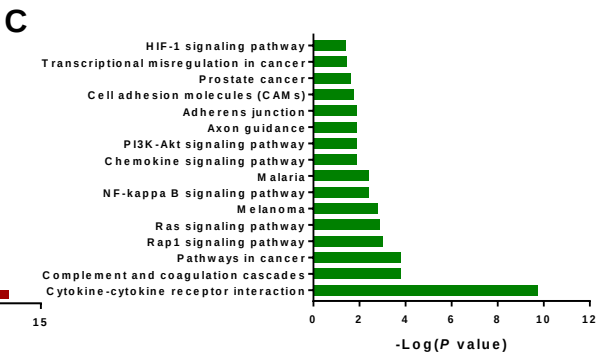



A
HC vs AA before IST
Proteins higher in AA
Serum
Plasma
478
27
16
B
AA5
AA7
AA7
AA3
AA2
AA1
AA1
AA3
HC-pool
AA6
AA4
AA5
AA2
AA4
AA6
C4A C4B
PAK6
F5
APOE
APOE3
APOE2
GFRA1
TFP1
GDF2
ACP5
COLEC11
IGFBP5
IL17RC
ITIH4
F3
SPON1
CHST15
SFRP1
IL18BP
TNC
AGRP
C5a
IGFBP4
HAMP
LBP
MRC1
C3b
-2
-1
0
1
2
C
C4A C4B
ALC
F5
PLT
PAK6
Hb
AMC
AMC/ALC
ARC
WBC
ANC
TFP1
GFRA1
ACP5
GDF2
APOE
APOE3
APOE2
COLEC11
IGFBP5
IL17RC
C5a
IGFBP4
HAMP
LBP
MRC1
C3b
ITIH4
F3
SFRP1
CHST15
SPON1
IL18BP
TNC
AGRP
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1

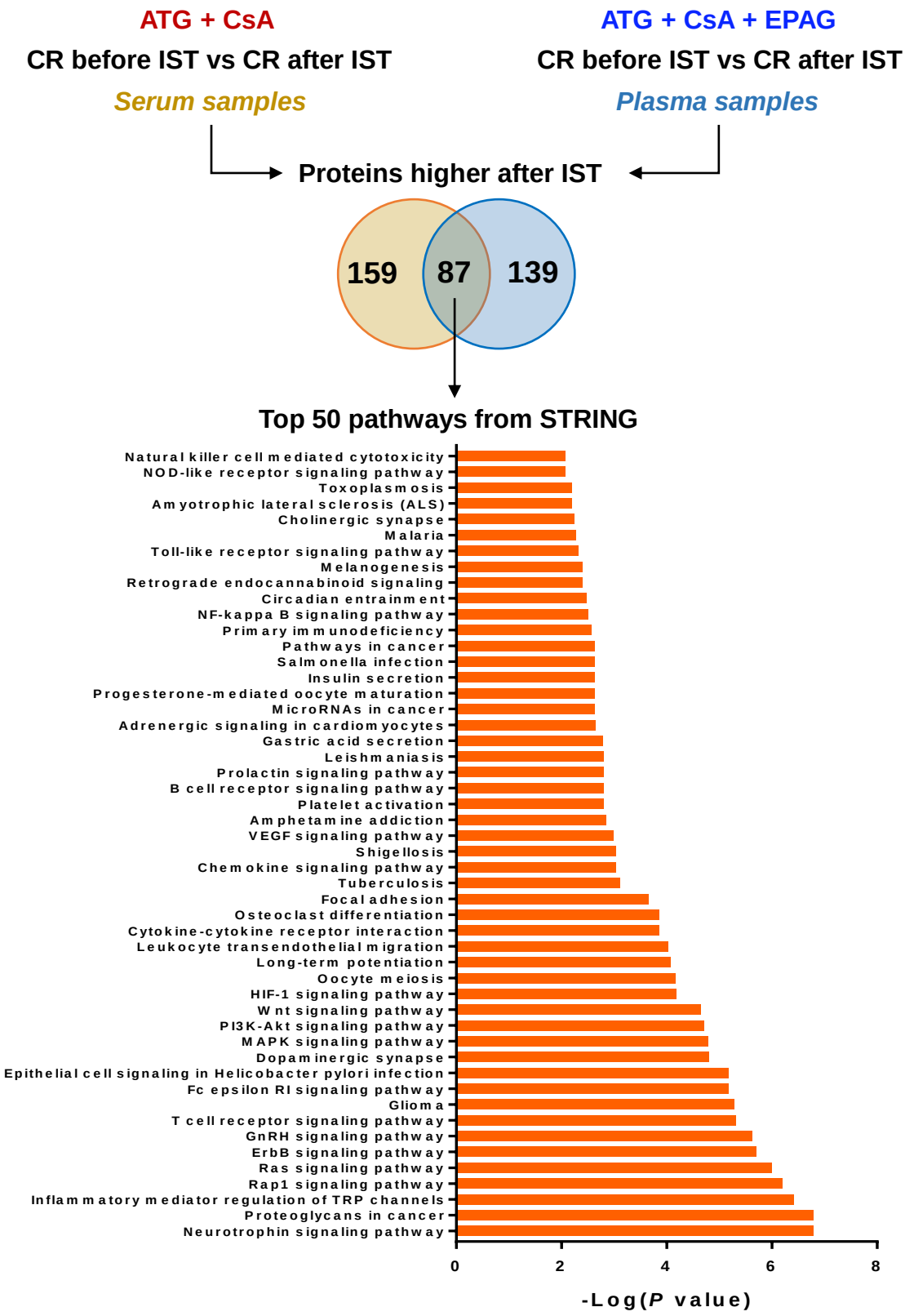
Supplementary Figure E5. Giudice V. et al.



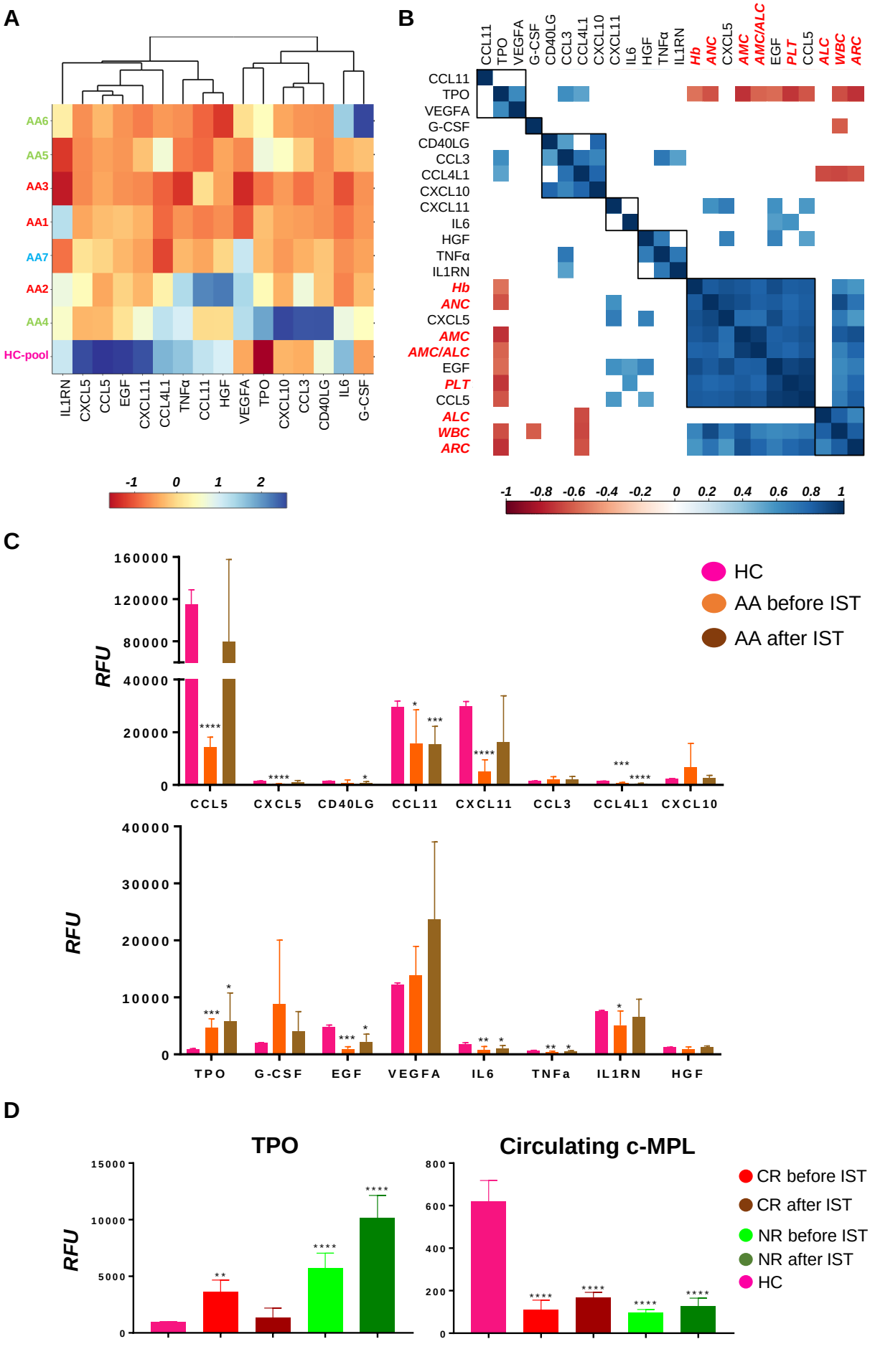




Supplementary Figure E6. Giudice V. et al.



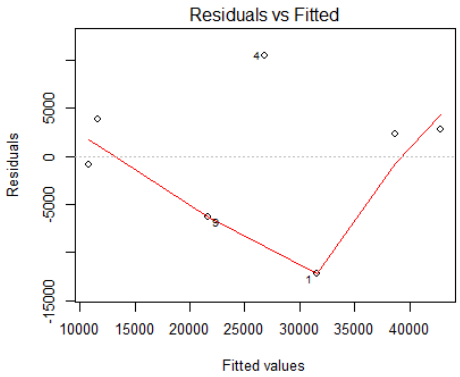
Supplementary Figure E7. Giudice V. et al.

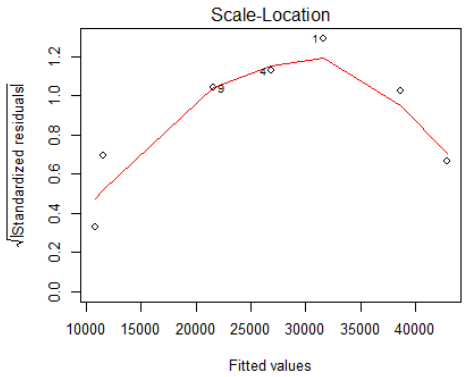


Supplementary Figure E8. Giudice V. et al.

A

Multiple linear regression
Disease severity vs DKK1 + CCL17 + SELL + HGF plasma levels

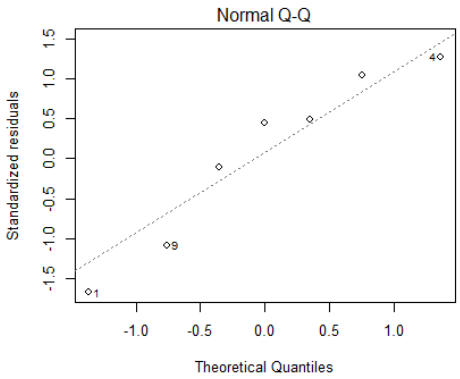


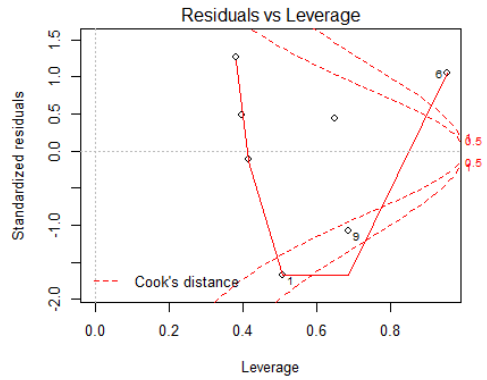



```
Analysis of Variance Table

Response: CCL17 + DKK1 + HGF + SELL
           Df    Sum Sq   Mean Sq F value  Pr(>F)
ANC         1 723038076 723038076  6.6287 0.08216 .
Plt         1 171440084 171440084  1.5717 0.29876
ARC         1  40091897  40091897  0.3676 0.58714
Residuals   3 327229714 109076571
```

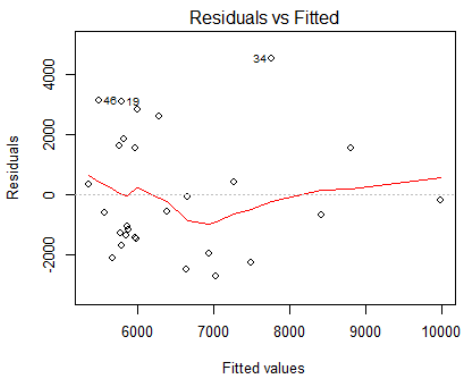
Multiple R^2 = 0.741
Adjusted R^2 = 0.481
F-statistic = 2.856 on 3 and 3 DF
p **= 0.206**

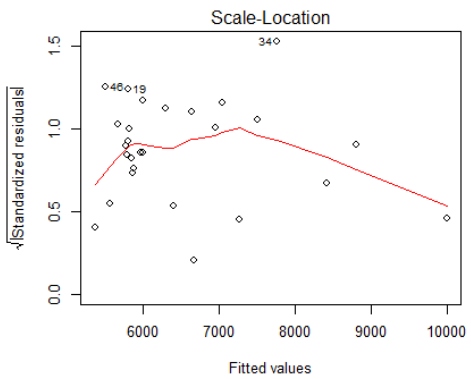




B

Multiple linear regression
Disease severity vs DKK1 + CCL17 + SELL + HGF serum levels






```
Analysis of Variance Table

Response: CCL17 + DKK1 + HGF + SELL
           Df    Sum Sq  Mean Sq F value  Pr(>F)
ANC         1  10623893 10623893  2.3812 0.13588
Plts        1  22292814 22292814  4.9967 0.03496 *
ARC         1    732420   732420  0.1642 0.68894
Residuals  24 107075693  4461487
```

Multiple R^2 = 0.239
Adjusted R^2 = 0.144
F-statistic = 2.514 on 3 and 24 DF
p **= 0.083**